

Lecture 22

Concentration Inequalities

Motivating Example

Opinion Poll: You want to know how many Americans support increasing taxes.

How to estimate this number?

Method: Pick at random n people from the population and ask for their opinion.

Want: The fraction of people supporting taxes increase among your poll is a "good" estimate to the overall fraction in the entire population.

a Mathy Formulation of the Problem

Suppose you have a coin w. heads prob. p

but you don't know p .

Goal: estimate p .

Method: Toss the coin n times

Let \hat{p} = frac. of times the coin turned out heads.

Want: $|p - \hat{p}| \leq \epsilon$ Can you guarantee it? **No!**

Settle for: $|p - \hat{p}| \leq \epsilon$ w. confidence $1 - \delta$.

$$\Pr[|p - \hat{p}| \leq \epsilon] \geq 1 - \delta$$

a Mathy Formulation of the Problem

Suppose you have a coin w. heads prob. p

but you don't know p .

Goal: estimate p .

Method: Toss the coin n times

Let \hat{p} = frac. of times the coin turned out heads.

Q: How large should n be to guarantee

$$\Pr[|p - \hat{p}| \leq \varepsilon] \geq 1 - \delta$$

A: $n \geq \frac{1}{4\varepsilon^2\delta}$ suffices.

For Example:

$$\varepsilon = 0.05 \quad \delta = 0.05$$

$$n = \frac{1}{4 \cdot (0.05)^3} = 2000 \text{ suffices.}$$

Q: Is it possible that everyone is above average (on the midterm)?

Q: How many students can be more than twice the average?

This is just Markov's Inequality.

Theorem (Markov's Inequality):

If X is a non-negative random variable then

for any constant $c \geq 0$

$$\Pr[X \geq c] \leq \frac{E[X]}{c}$$

Another formulation: for any constant $k \geq 0$

$$\Pr[X \geq k \cdot E[X]] \leq \boxed{}$$

Theorem (Markov's Inequality):

If X is a non-negative random variable then

for any constant $c \geq 0$

$$\Pr[X \geq c] \leq \frac{\mathbb{E}[X]}{c}$$

Proof:

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) \cdot \Pr[\omega]$$

$$\geq \sum_{\omega: X(\omega) \geq c} X(\omega) \cdot \Pr[\omega]$$

$$\geq \sum_{\omega: X(\omega) \geq c} c \cdot \Pr[\omega]$$

$$= c \cdot \sum_{\omega: X(\omega) \geq c} \Pr[\omega] = c \cdot \Pr[X \geq c] \quad \blacksquare$$

Theorem (Markov's Inequality):

If X is a non-negative random variable then

for any constant $c \geq 0$

$$\Pr[X \geq c] \leq \frac{E[X]}{c}$$

Example: Coin Tosses

You toss a fair coin n times. X - number of heads.

What's the prob. of $X \geq \frac{3}{4}n$?

using Markov's Ineq. $\Pr[X \geq \frac{3}{4}n] \leq \frac{E[X]}{\frac{3}{4}n} = \frac{\frac{1}{2}n}{\frac{3}{4}n} = \frac{2}{3}$

Is this a good upper bound?

Theorem (Markov's Inequality):

If X is a non-negative random variable then

for any constant $k \geq 0$

$$\Pr[X \geq k \cdot \mathbb{E}[X]] \leq \frac{1}{k}$$

Example 2: Let $\mu \geq 0, k \geq 1$.

$$X = \begin{cases} \mu^k, & \text{w.p. } \frac{1}{k} \\ 0, & \text{otherwise} \end{cases}$$

$$\mathbb{E}X = \boxed{}$$

$$\Pr[X \geq k \cdot \mathbb{E}X] = \boxed{}$$

Theorem (Markov's Inequality):

If X is a non-negative random variable then

for any constant $c \geq 0$

$$\Pr[X \geq c] \leq \frac{\mathbb{E}[X]}{c}$$

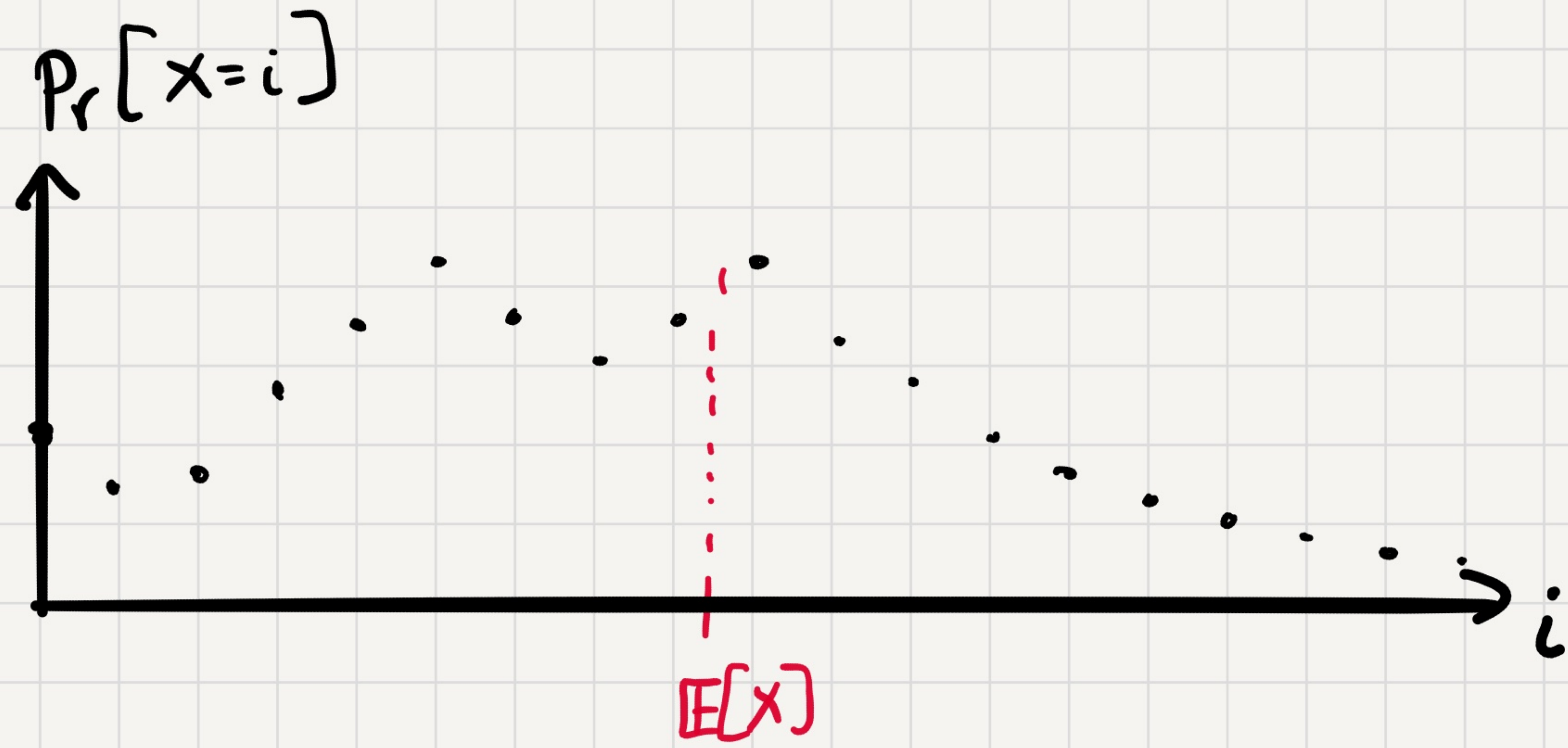
Q: What happens if X can be negative?

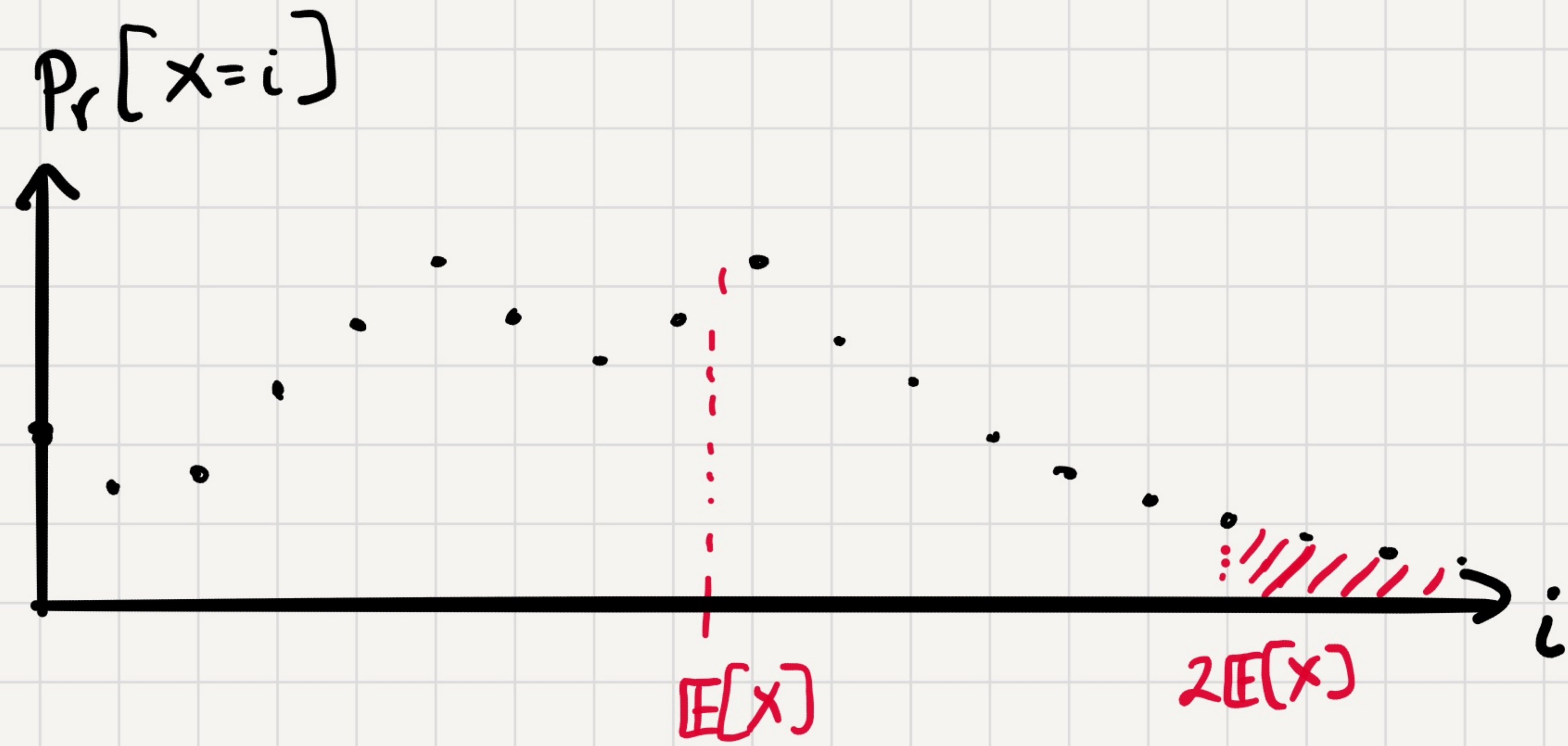
Example:

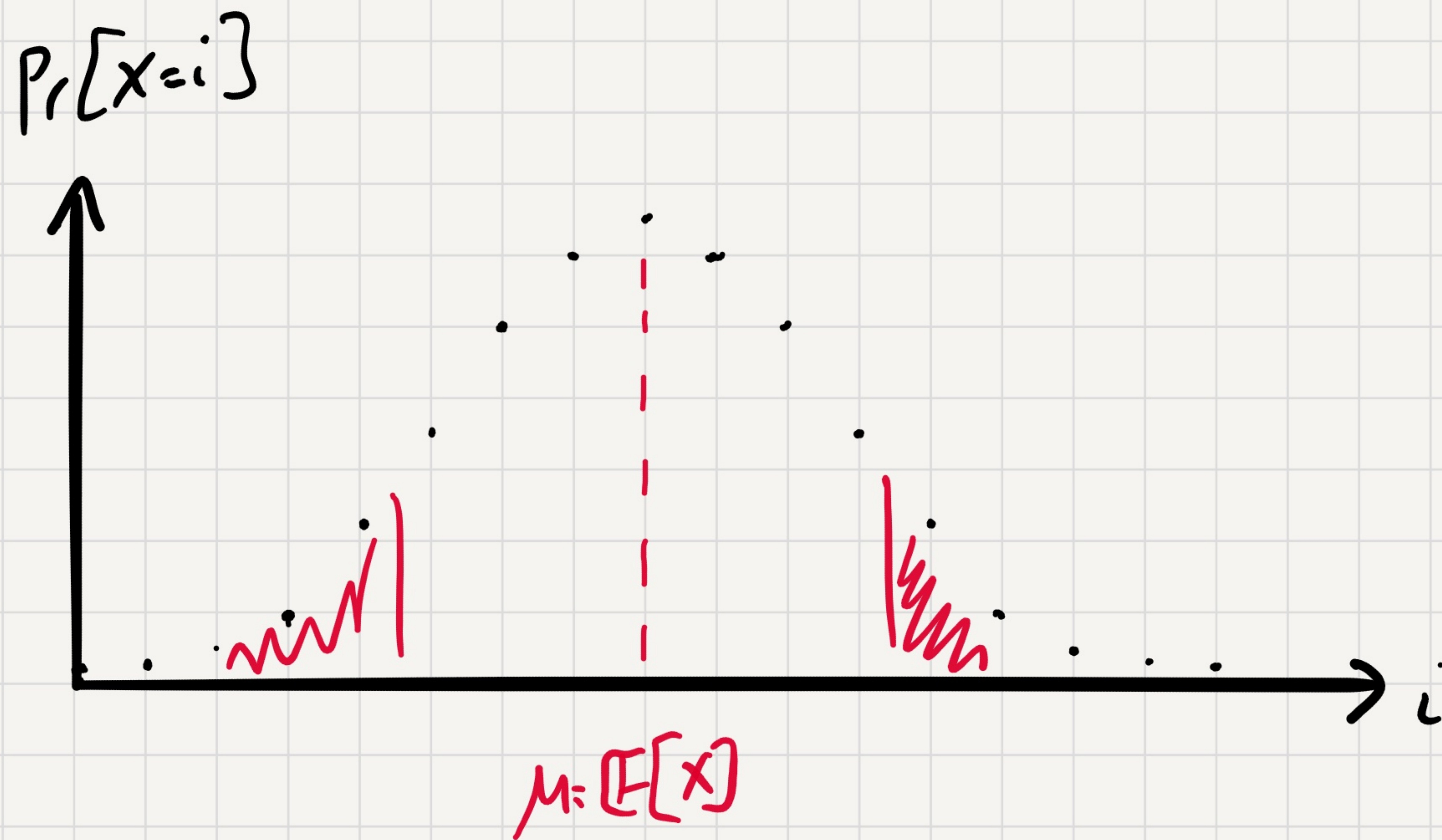
$$X = \begin{cases} +1, & \text{w.p. } \frac{1}{2} \\ -1, & \text{w.p. } \frac{1}{2} \end{cases}$$

$$\mathbb{E}X = 0$$

$$\Pr[X \geq 1] = \frac{1}{2} \neq \frac{\mathbb{E}X}{1}$$







What's the prob. to be "far" from the mean?

That is captured by Chebyshev's Inequality.

Theorem (Chebyshev's Inequality)

Let X be a r.v. with expectation μ

Then
$$\Pr [|X - \mu| \geq c] \leq \frac{\text{Var}(X)}{c^2}$$

for any constant $c \geq 0$.

Equivalent Formulation:

If X has mean μ and stdev σ then

for any $k \geq 0$
$$\Pr [|X - \mu| \geq k\sigma] \leq$$



Theorem (Chebyshev's Inequality)

Let X be a r.v. with expectation μ

Then
$$\Pr[|X - \mu| \geq c] \leq \frac{\text{Var}(X)}{c^2}$$

for any constant $c \geq 0$.

Proof:
$$\Pr[|X - \mu| \geq c] = \Pr[(X - \mu)^2 \geq c^2]$$

Let $Y = (X - \mu)^2$. Y is a non-negative r.v.

$$E[Y] = E[(X - \mu)^2] = \text{Var}(X).$$

$$\Pr[|X - \mu| \geq c] = \Pr[Y \geq c^2] \leq \frac{E[Y]}{c^2} = \frac{\text{Var}(X)}{c^2} \quad \blacksquare$$

Theorem (Chebyshev's Inequality)

Let X be a r.v. with expectation μ

$$\text{Then } \Pr[|X - \mu| \geq c] \leq \frac{\text{Var}(X)}{c^2}$$

for any constant $c \geq 0$.

Example: Coin Tosses

You toss a fair coin n times. X - number of heads.

What's the prob. of $X \geq \frac{3}{4}n$?

$$X \sim \text{Bin}(n, \frac{1}{2}). \quad E[X] = n/2 \quad \text{Var}(X) = n \cdot p \cdot (1-p) = \frac{n}{4}$$

$$\Pr[X \geq \frac{3}{4}n] \leq \Pr[|X - \frac{n}{2}| \geq \frac{1}{4}n] \leq \frac{\text{Var}(X)}{(\frac{1}{4}n)^2} = \frac{n/4}{(n/4)^2} = \frac{4}{n}$$

goes to 0 as $n \rightarrow \infty$.

Back to Motivating Example

You have a coin w. unknown heads prob. p .

Want to estimate p .

Method: toss the coin n times

X = number of heads

$\hat{p} = \frac{X}{n}$ "empirical mean"

$$\Pr[|p - \hat{p}| \geq \varepsilon] \leq ?$$

Back to Motivating Example

You have a coin w. unknown heads prob. p .

Toss the coin n times.

X = number of heads. $X = X_1 + \dots + X_n$

X_i - indicator that i^{th} coin is heads

$$\hat{p} = \frac{X}{n} = \frac{X_1 + \dots + X_n}{n}$$

$$E[\hat{p}] = E\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{n p}{n} = p.$$

$$\text{Var}[\hat{p}] = \frac{\text{Var}(X)}{n^2} = \frac{\text{Var}(X_1) + \dots + \text{Var}(X_n)}{n^2} = \frac{n p \cdot (1-p)}{n^2} \leq \frac{1/4}{n}$$

$$\text{Pr}[|\hat{p} - p| \geq \varepsilon] \leq \frac{\text{Var}[\hat{p}]}{\varepsilon^2} = \frac{1}{4n\varepsilon^2}$$

Back to Motivating Example

You have a coin w. unknown heads prob. p .

Toss the coin n times.

X = number of heads. $X = X_1 + \dots + X_n$

X_i - indicator that i th coin is heads

$$\hat{p} = \frac{X}{n} = \frac{X_1 + \dots + X_n}{n}$$

$$E[\hat{p}] = p$$

$$\text{Var}[\hat{p}] = \frac{p(1-p)}{n} \leq \frac{1}{4n}$$

$$\Pr[|\hat{p} - p| \geq \varepsilon] \leq \frac{\text{Var}[\hat{p}]}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$$

Theorem: If $n \geq \frac{1}{4\varepsilon^2\delta}$ then

$$\Pr[|\hat{p} - p| \geq \varepsilon] \leq \delta.$$

Weak Law of Large Numbers

If X_1, X_2, X_3, \dots are independent identically distributed (i.i.d.)

with $E X_i = \mu$, $\text{Var}(X_i) < \infty$, then for every $\varepsilon > 0$,

$$\Pr \left[\left| \frac{X_1 + \dots + X_n}{n} - \mu \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0.$$

Proof:

Let $\sigma^2 = \text{Var}(X_i)$

$$S_n = X_1 + \dots + X_n$$

$$Y_n = \frac{S_n}{n}$$

$$E[Y_n] = \boxed{}$$

$$\text{Var}[Y_n] = \boxed{}$$

$$\Pr \left[|Y_n - \mu| \geq \varepsilon \right] \leq$$

Q: Is Chebyshev's Inequality best possible?

A: Yes, if we only know μ and σ .

For example: For any μ, σ and any $k \geq 1$

$$X = \begin{cases} \mu + \sigma k & \text{w.p. } \frac{1}{2k^2} \\ \mu - \sigma k & \text{w.p. } \frac{1}{2k^2} \\ \mu & \text{w.p. } 1 - \frac{1}{k^2} \end{cases}$$

$$E[X] =$$

$$\text{Var}[X] =$$

$$\Pr[|X - E[X]| \geq \sigma k] =$$

Theorem:

$$\text{If } n \geq \frac{1}{4\epsilon^2\delta}$$

$$\Pr(|\hat{P} - P| \geq \epsilon) \leq \delta.$$

Q: Is $\frac{1}{4\epsilon^2\delta}$ best possible?

A: No. $n = \frac{1}{2\epsilon^2} \ln\left(\frac{2}{\delta}\right)$ suffices

using Chernoff bounds - out of scope.

Confidence is "cheap"

Accuracy is "expensive."

Confidence Interval

You don't know the prob. p of heads.

You tossed it a 1,000,000 times and get H 35% of the times.

How confident are you that $p = 0.35$?

" " " " that $p \in [0.3, 0.4]$?

Def'n: An interval $[a, b]$ is a 95% - confidence interval for an unknown quantity p if

$$\Pr [p \in [a, b]] \geq 0.95$$

The interval $[a, b]$ is calculated based on an experiment.

Confidence Interval

You don't know the prob. p of heads.

Theorem: \hat{p} = frac. of heads in n tosses

$$\Pr [|\hat{p} - p| \leq \varepsilon] \geq 1 - \frac{1}{4n\varepsilon^2}$$

We can pick $n = \frac{1}{4\varepsilon^2 \cdot 0.05} = \frac{5}{\varepsilon^2}$ to get confidence 95%.

What's the confidence interval?

$$a = \hat{p} - \varepsilon \quad b = \hat{p} + \varepsilon$$

$$\Pr [p \in [a, b]] = \Pr [|p - \hat{p}| \leq \varepsilon] \geq 0.95.$$